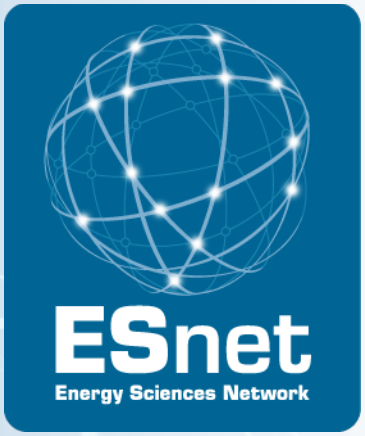


Network Engineering Group Update

Joe Metzger, Network Engineer

ESnet Network Engineering Group

ESCC

January 17 2013

Outline

Current Network Status

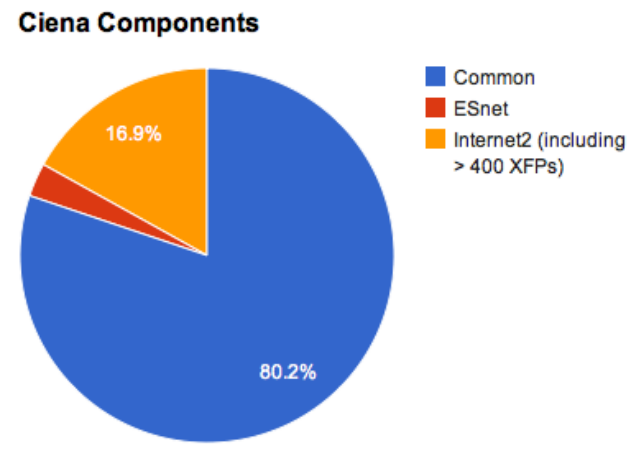
Short-term next steps (6 months)

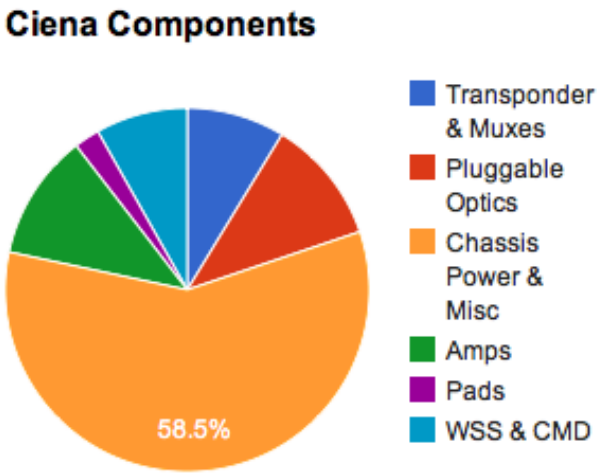
What comes after that?

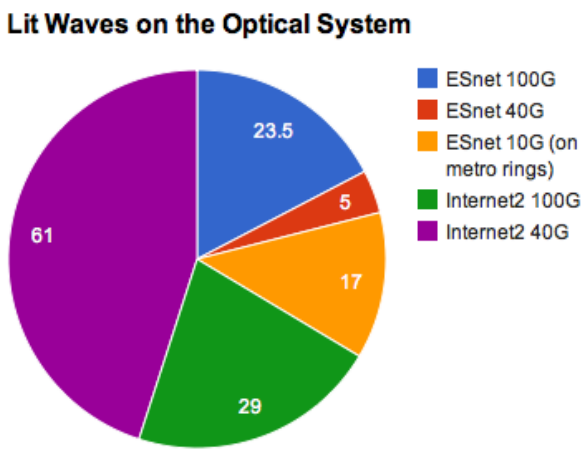
Current Status ESnet5 Transport Network

Shared Optical System

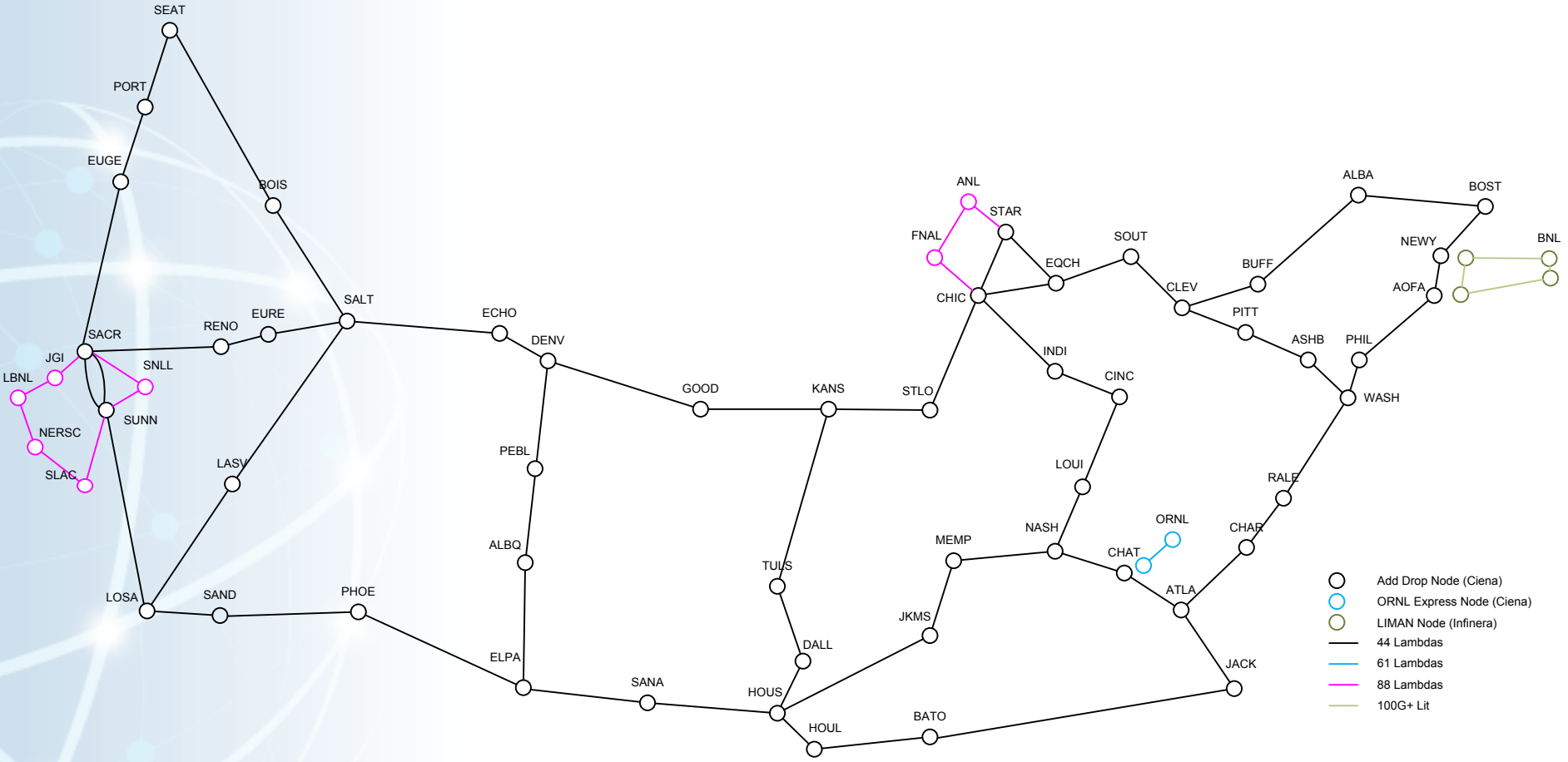
- All common components deployed & commissioned
- Dec 4th Optical System Inventory report has 6495 components!
  - 341 nodes total, 60+ are add/drop/regen
  - 80% are common, the rest are dedicated to ESnet or Internet2 (over 500 are XFPs.)
  - Sunnyvale has 4 32-slot shelves
  - Sacramento is a 7-direction node
- Still refining processes for sharing spectrum, chassis, configuration, management, etc.



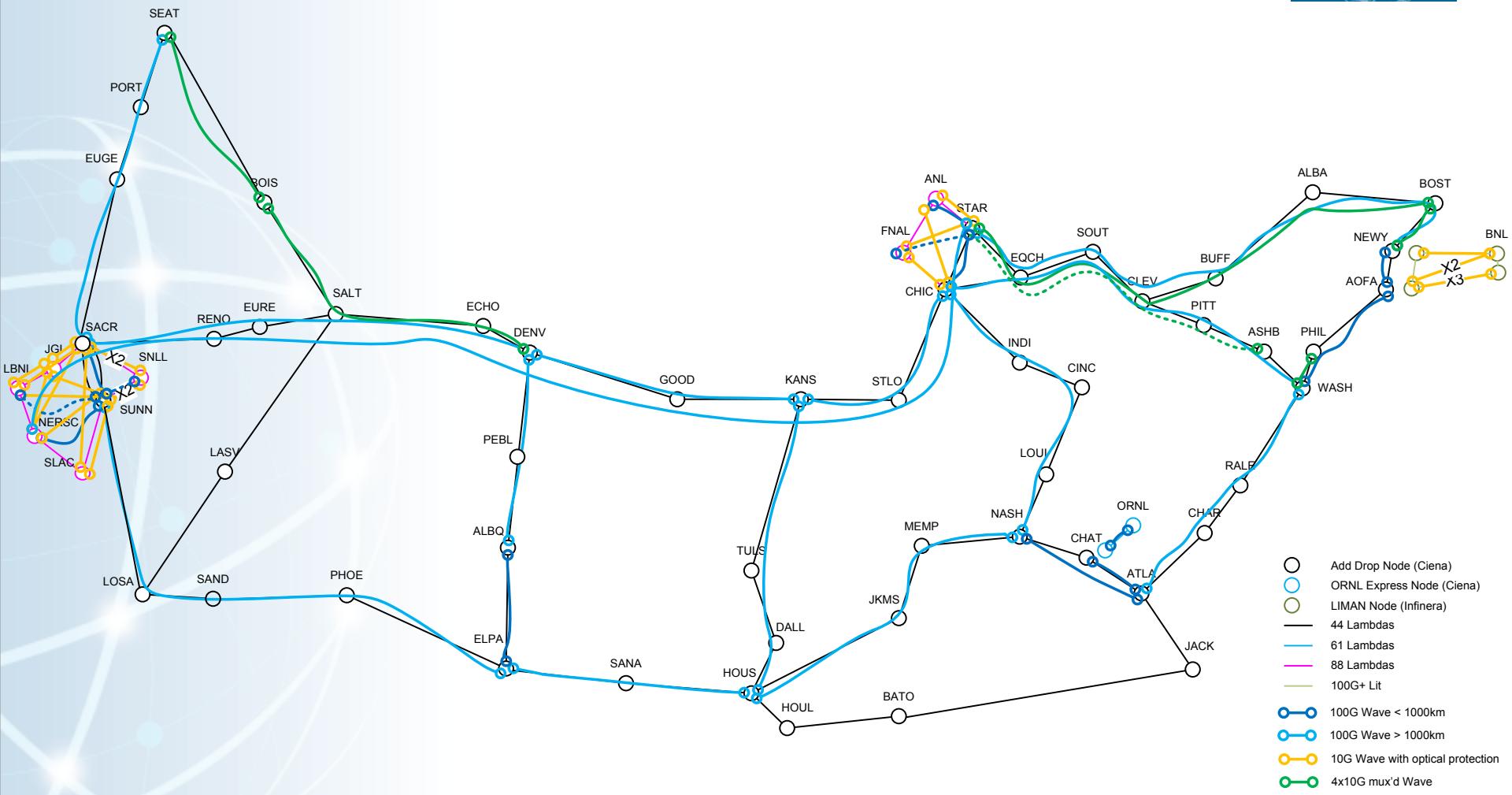




ESnet5 Optical System January 2012



ESnet5 Optical System January 2012
Lit Waves



ESnet5 Routed Network

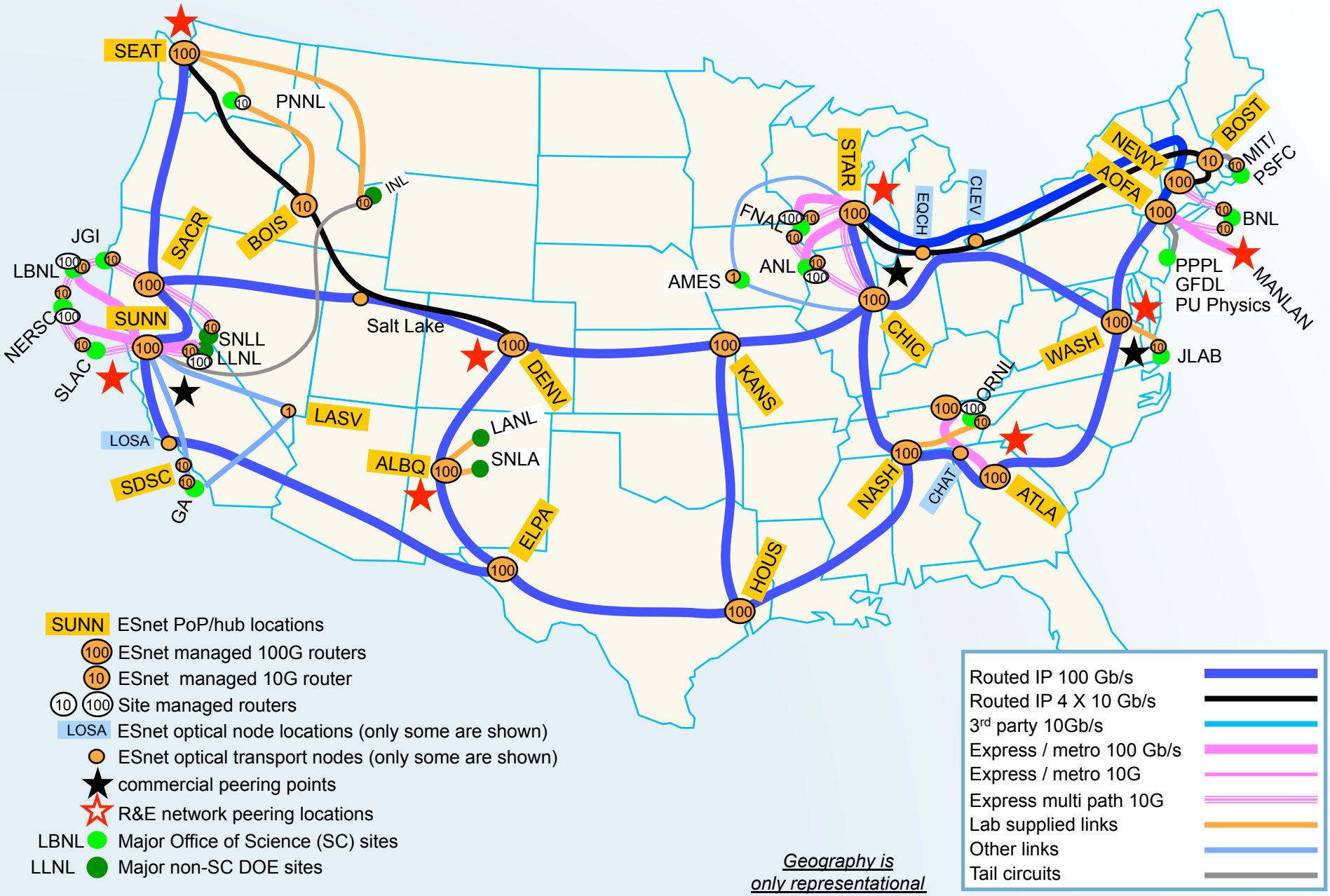
Routers

- 16 new Alcatel Lucent (ALU) 7750-SR12
  - 10-slot router with up to 200G per slot
  - 56 100G interfaces & 200+ 10G interfaces
- 35 existing Juniper MX's
  - Used in 10G hubs, commercial exchange points, sites
- 12 existing Juniper M7i & M10i
  - For terminating links slower than GE
- 5 very old Cisco 7206
  - For terminating links slower than GE

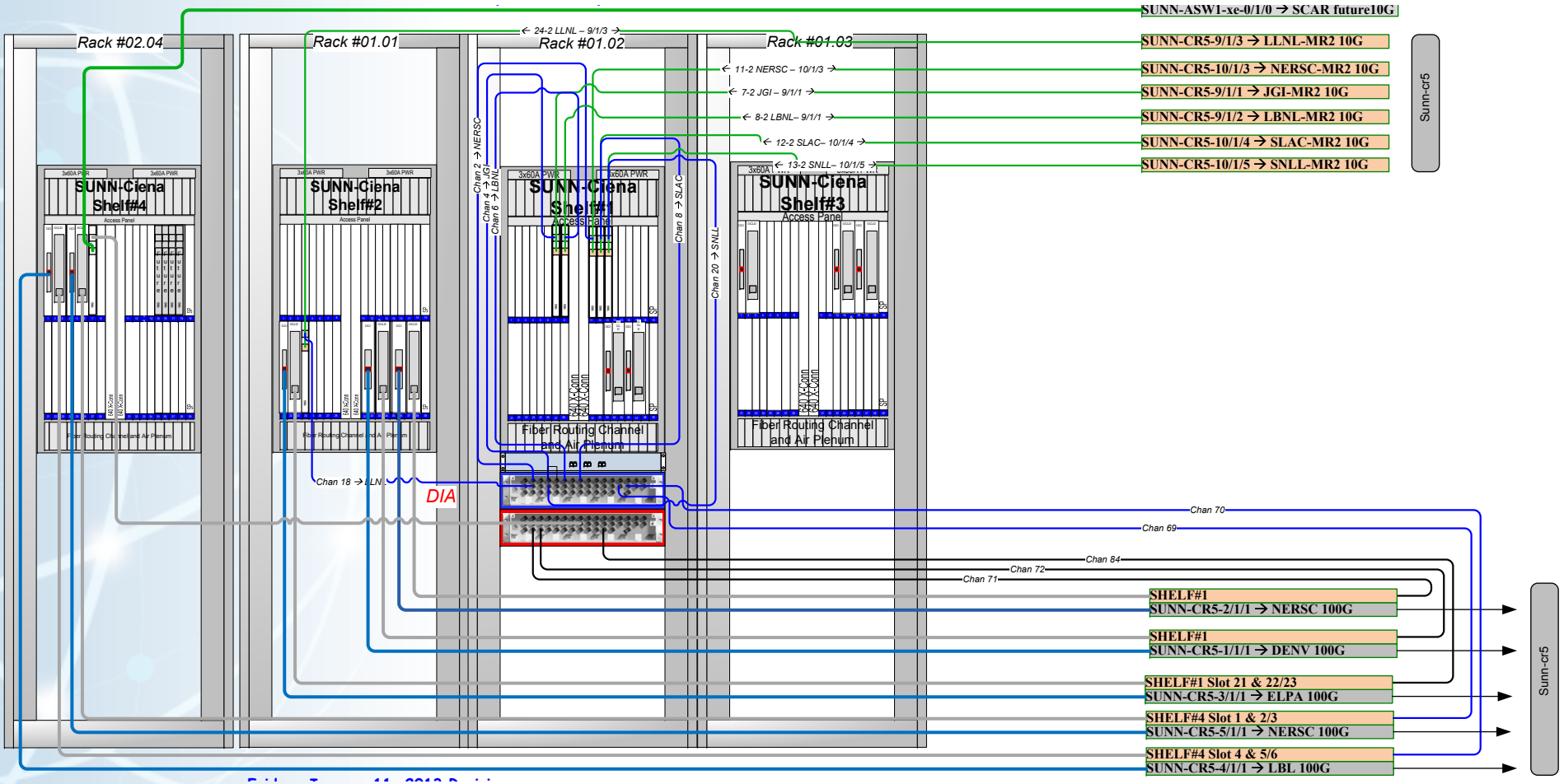
Services

- Standard routed IP (including full Internet services)
- Point to Point Dynamic Virtual Circuits using OSCARS
- Various overlay networks (Private VPN's, LHCONE VRF)

ESnet5 January 2013



Example: Optical System in Sunnyvale



Next 6 Months

- 100G production connections to:
  - ANL, BNL, FNAL, LBNL, LLNL, NERSC & ORNL
- 100G production connections to exchange points & R&E peers:
  - MANLAN, Starlight, PACWAVE, WIX, Internet2
- 40G into Equinix Ashburn & re-arranging our Washington DC ring to provide diverse backbone connections for JLAB & other sites in the area
- Diverse fiber laterals & diverse optical nodes at ANL & FNAL
- Deploying diverse routers at ORNL
- Lots of cleanup & consolidation at the hubs, moving connections from the MX's to the ALUs
- Normalize 100G Testbed infrastructure
- Swap out our un-supported 'third party' 10x10 MSA CFPs in our Ciena interfaces with Ciena supported ones covered by 4-hour on-site maintenance contract

Disclaimer for rest of this slide deck

- The rest of this slide deck makes all sorts of assumptions
  - Some have been carefully thought out
  - Some haven't had much thought at all
  - Some we don't even realize we are making
  - Some of them are sure to be violated

- It is intended to spur thought and discussion.
  - Not to come up with the right answers today, but to make sure we are not missing any important questions.

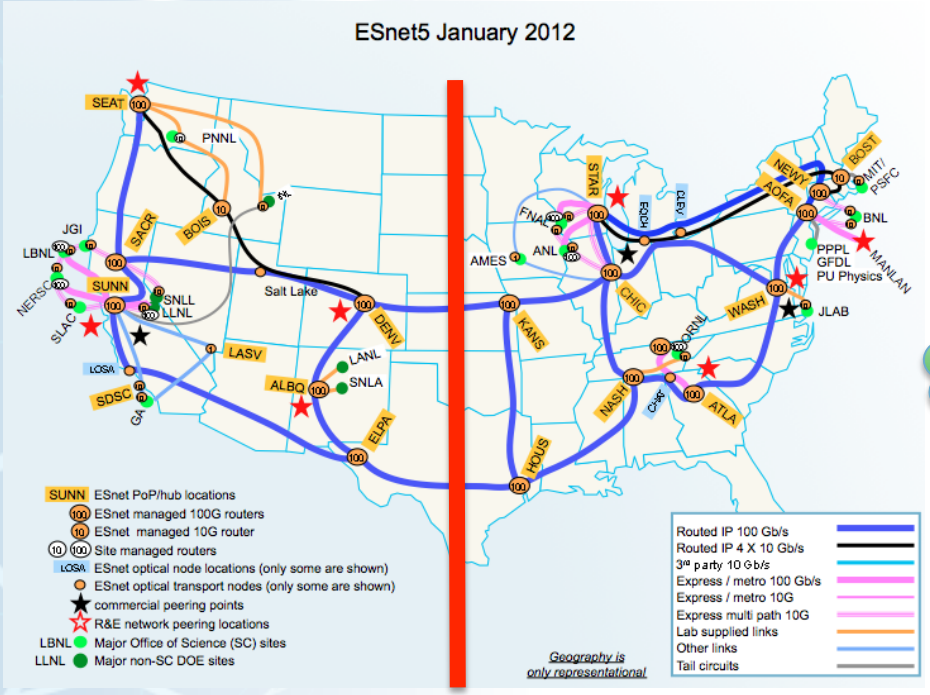
Backbone & Traffic Metrics

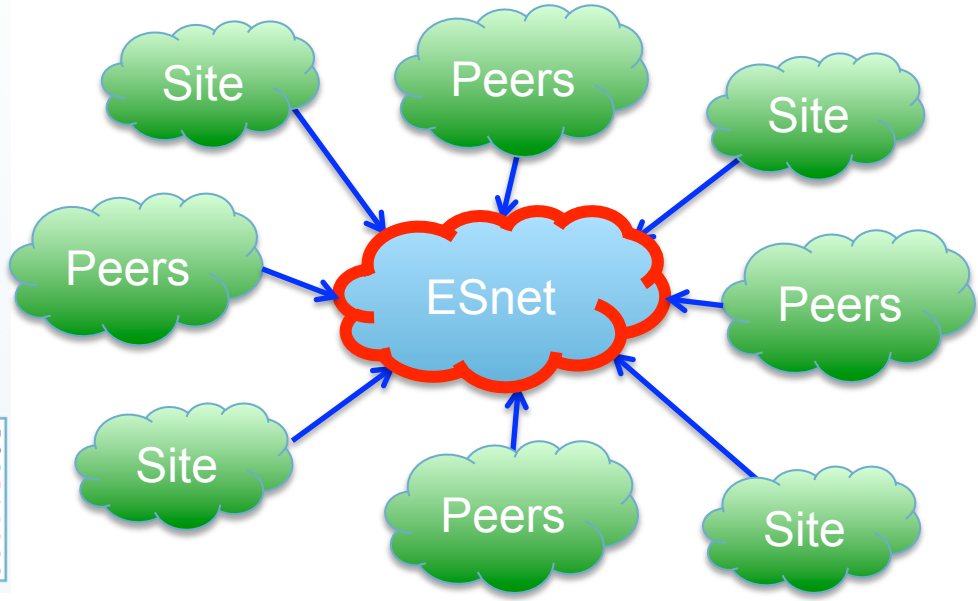
Backbone Bisection Bandwidth:

Sum of Links crossing the red line

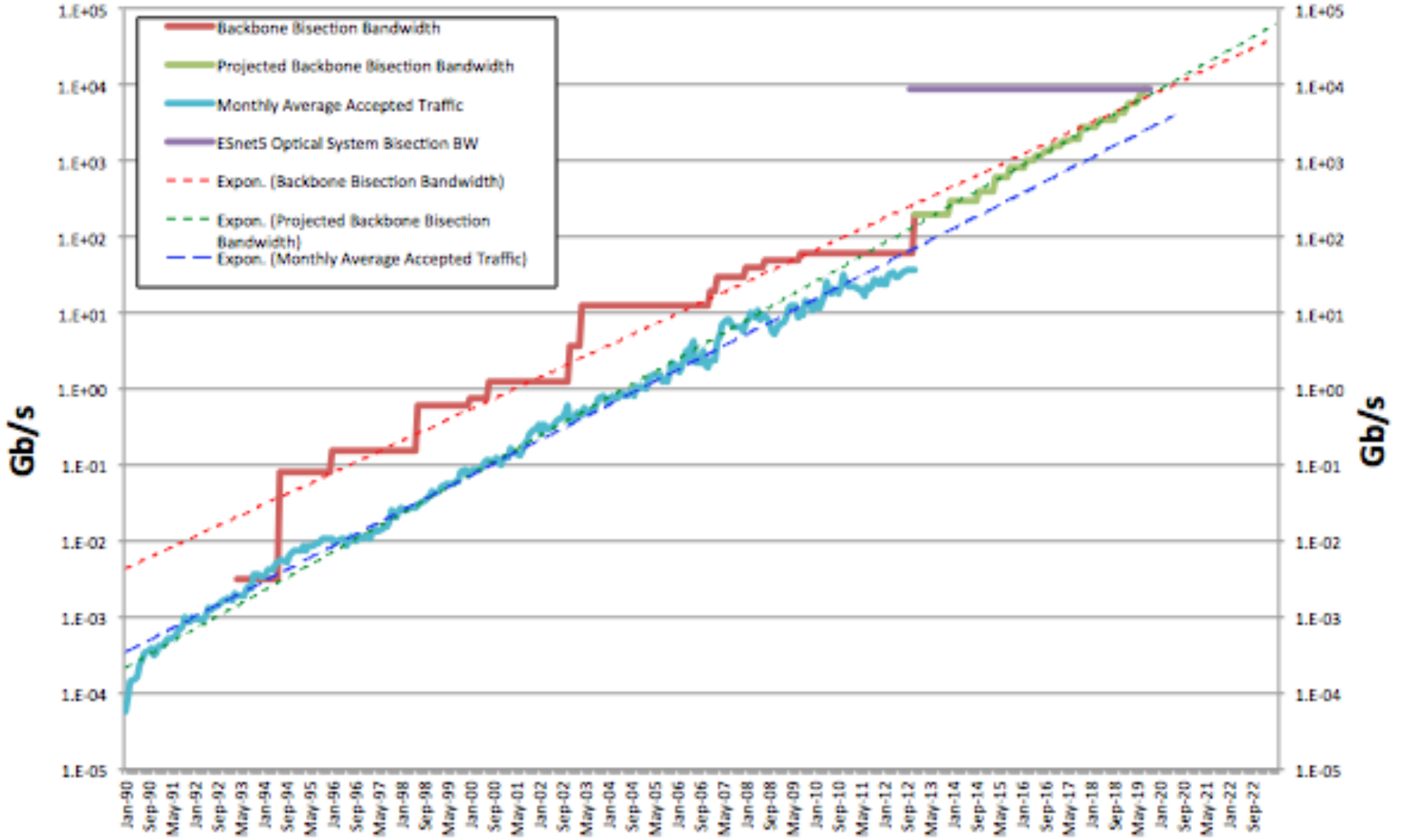
Average Accepted Traffic:

Sum of inbound traffic crossing red line from Sites & Peers, converted to Gb/s

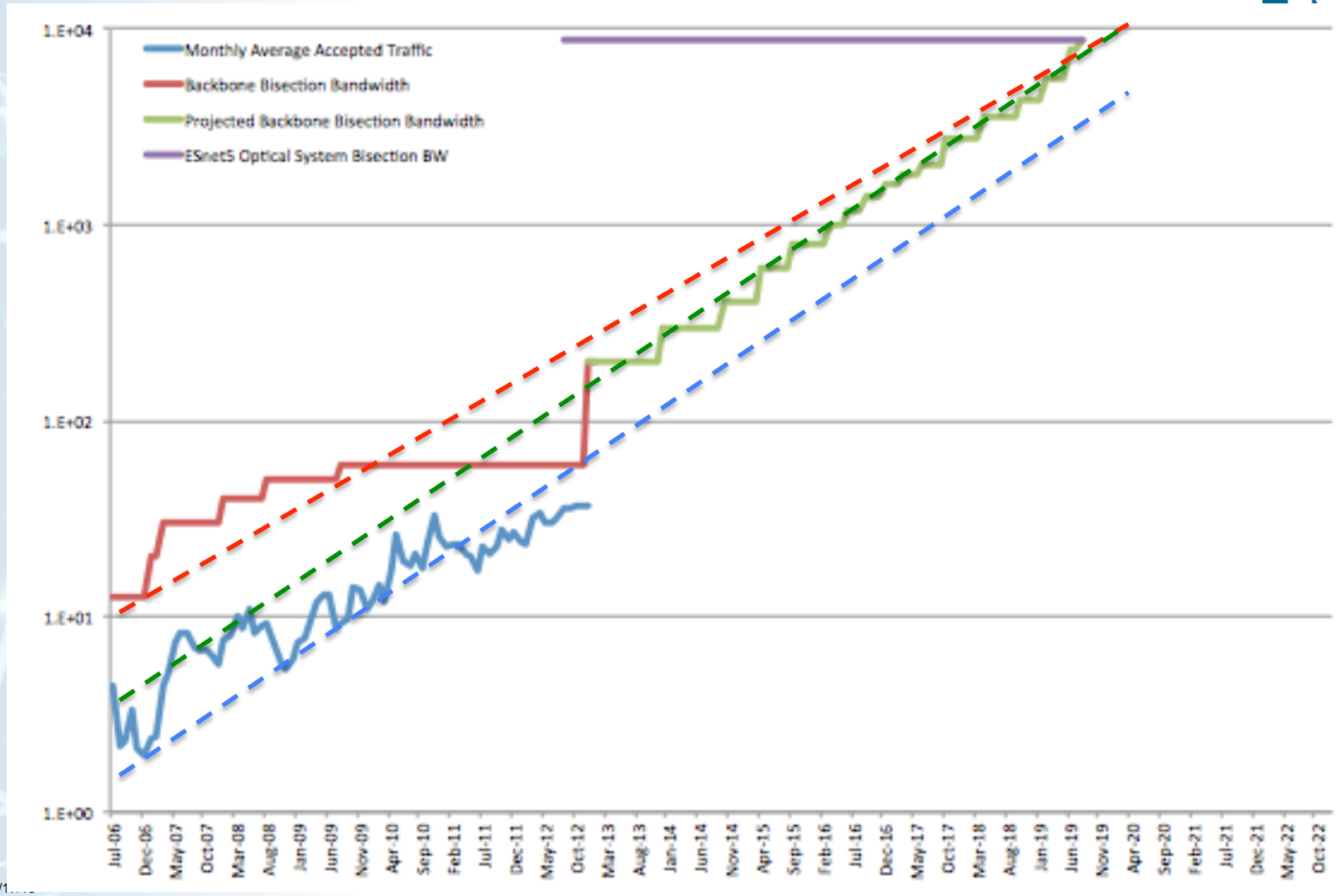




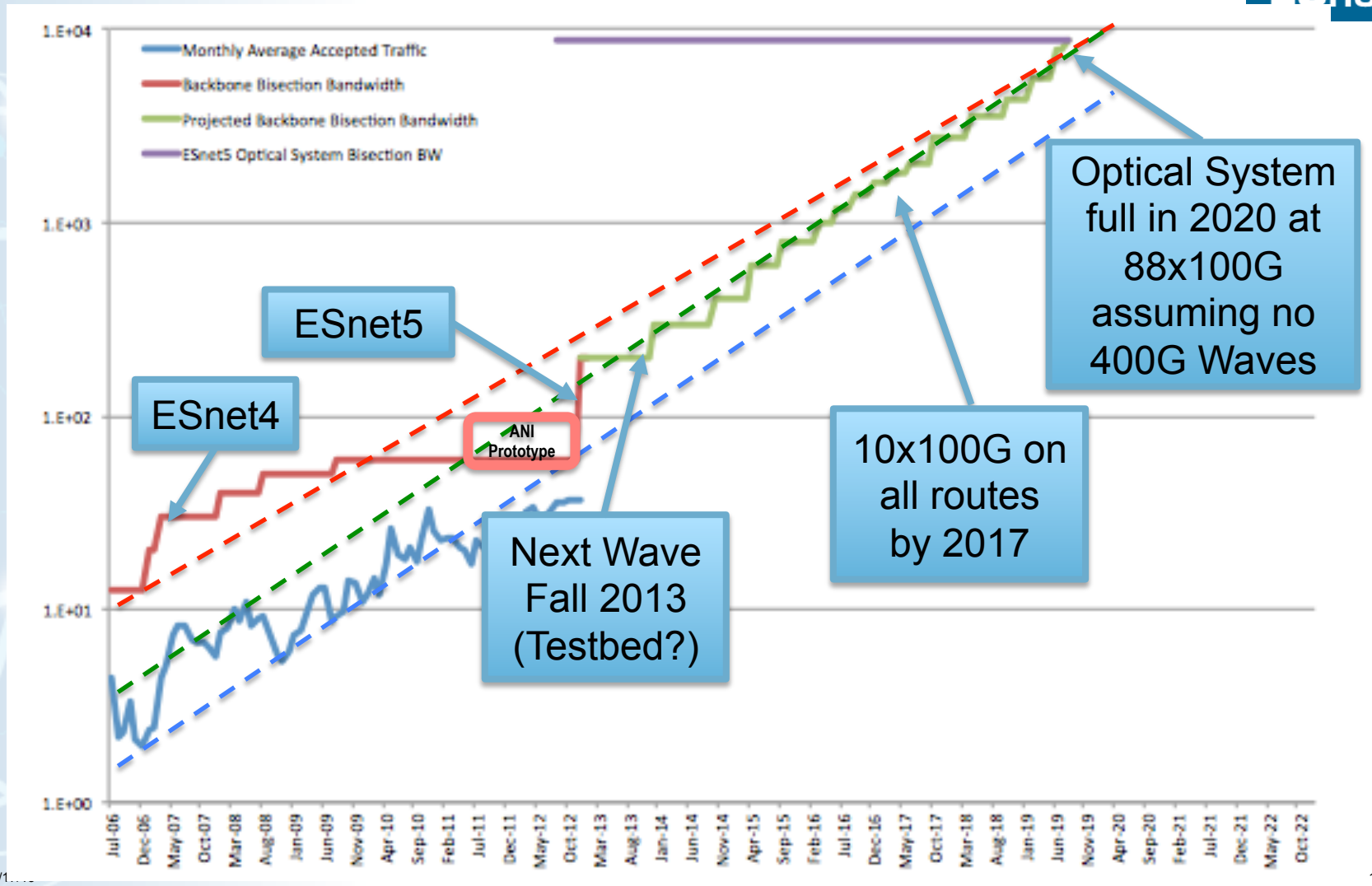
ESnet Traffic vs Backbone Capacity



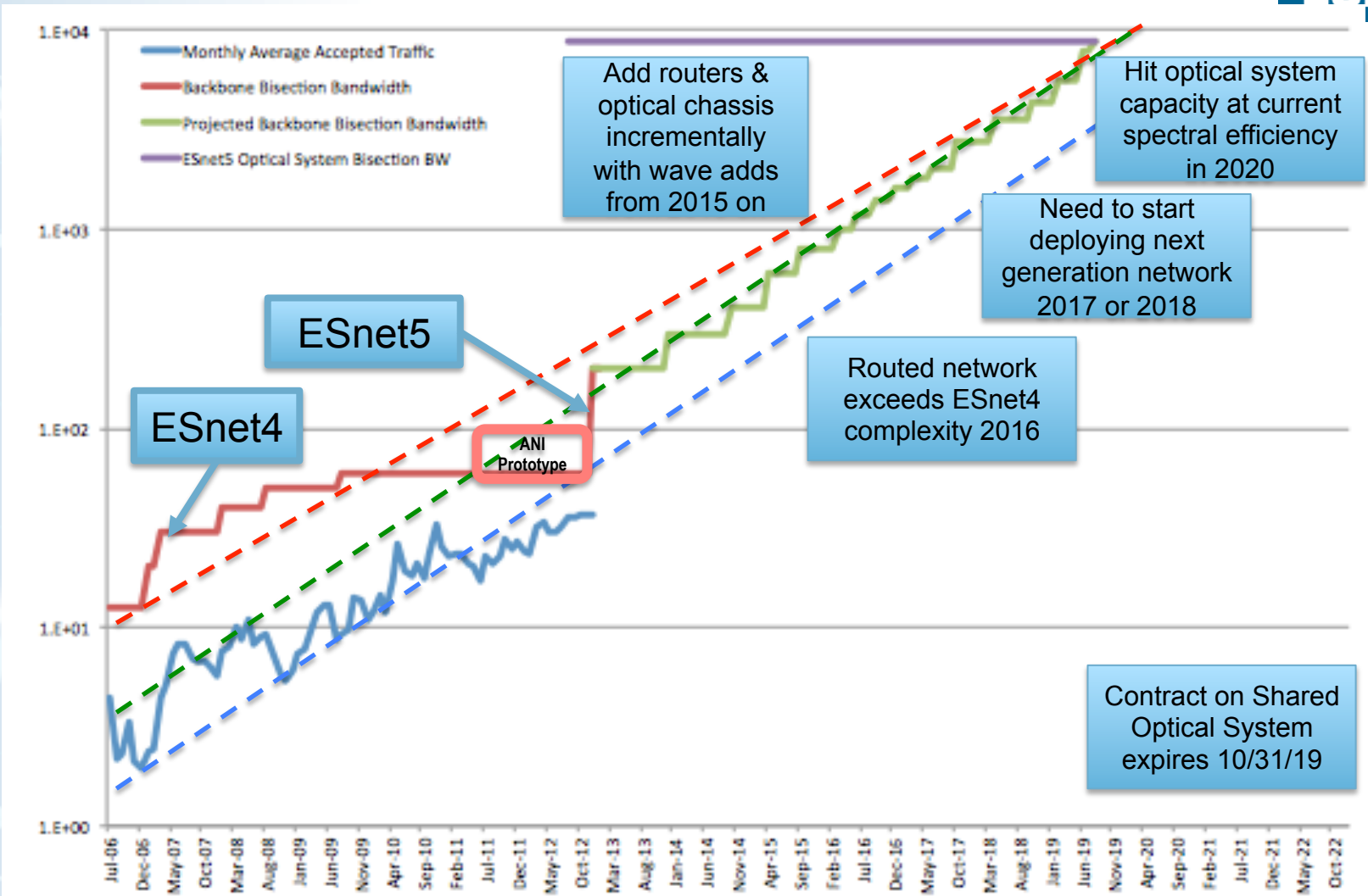
Projections for next 7 years



Projections for next 7 years



Cost Inflection Points next 7 years



Things to consider

- How much 'head-room' is enough
  - Need less 'bisection bandwidth' if traffic is local vs cross-country
  - Need more for unpredictable users
  - Need less for predictable steady-state users
    - Could automated science work-flow systems be more predictable, and sustain utilization >75%?

- Some changes that could drive growth faster than projected
  - Science DMZ adoption reducing impacts of poorly performing firewalls & security infrastructure
  - PerfSONAR adoption & deployment allowing people to discover & fix broken infrastructure
  - Data mobility tools enhancement & adoption
  - Commercial Adoption rate of 100GE will affect the price curve of 100G components driving up lab adoption rates
  - Changes in science data access models could affect growth rates in either direction

Evolution of ESnet5

- Simplifying Hubs where possible
  - Removing ‘extra’ points of failure
- Optical Express circuits
  - Some circuits could bypass all routers
- Expanding International reach, capacity & control

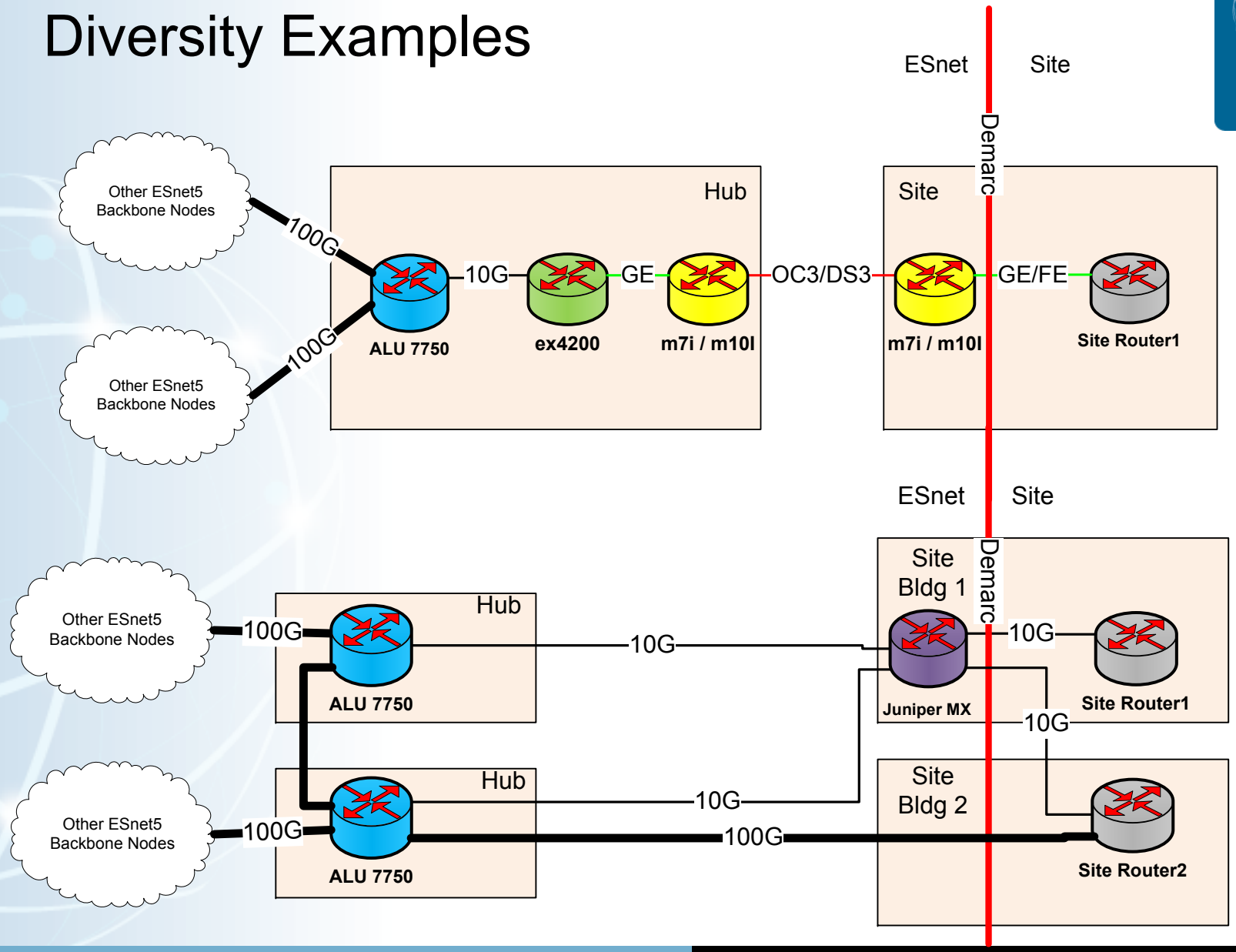
- For now, we have collapsed Nx10GE of IP & SDN traffic onto single 100G circuits using QOS for logical separation.
- We anticipate moving back to physical separation of traffic onto different links as additional circuits are added.

Evolution of Site Interconnections

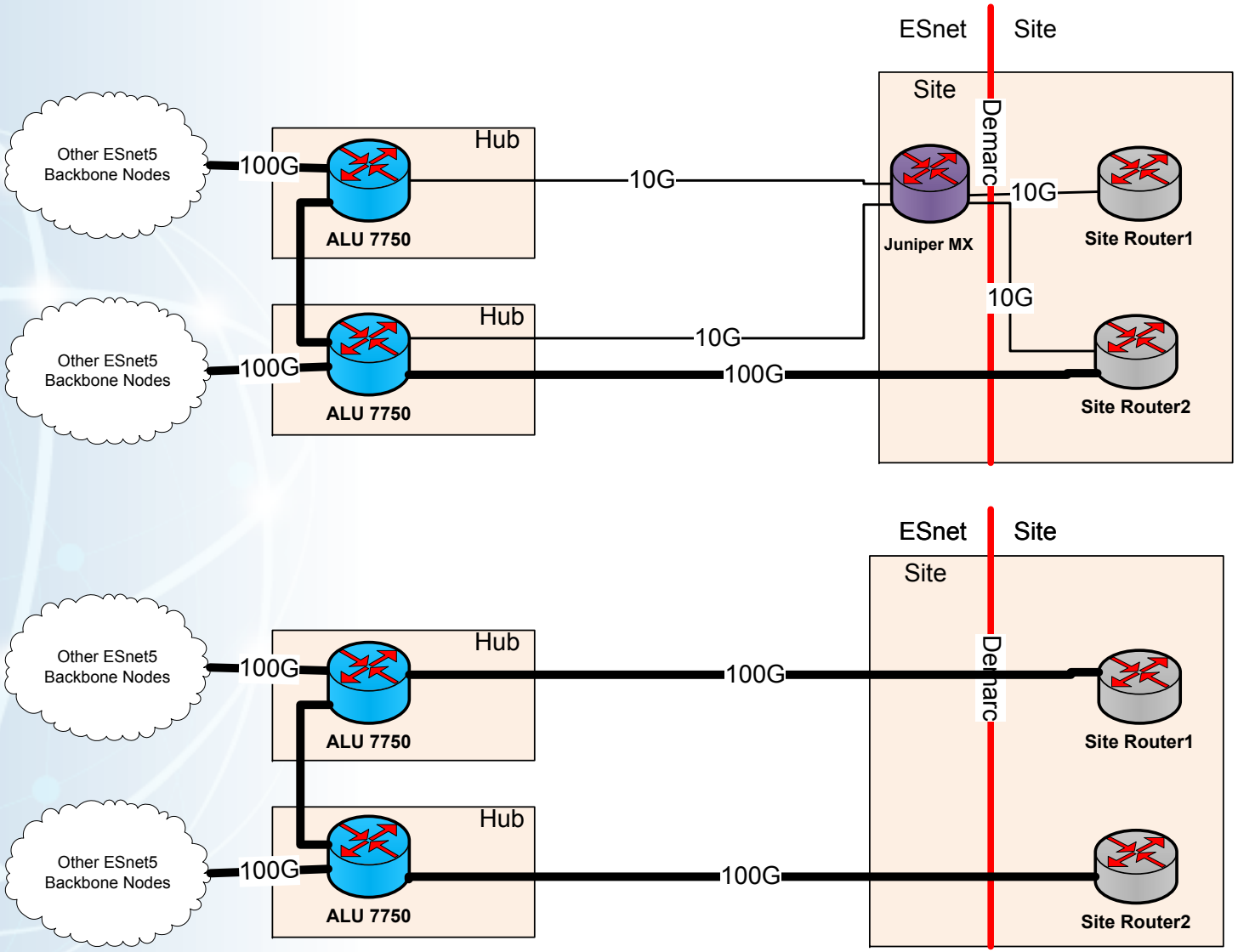
Goal: provide the best possible service within the constraints

- Dimensions of ‘best’ that are important:
  - Bandwidth
  - Reliability
  - Predictability
  - Risk tolerance
  - Debug-ability
  - Flexibility
  - Cost
- Discussion Question: What are the comparative weights of the different dimensions?

Diversity Examples

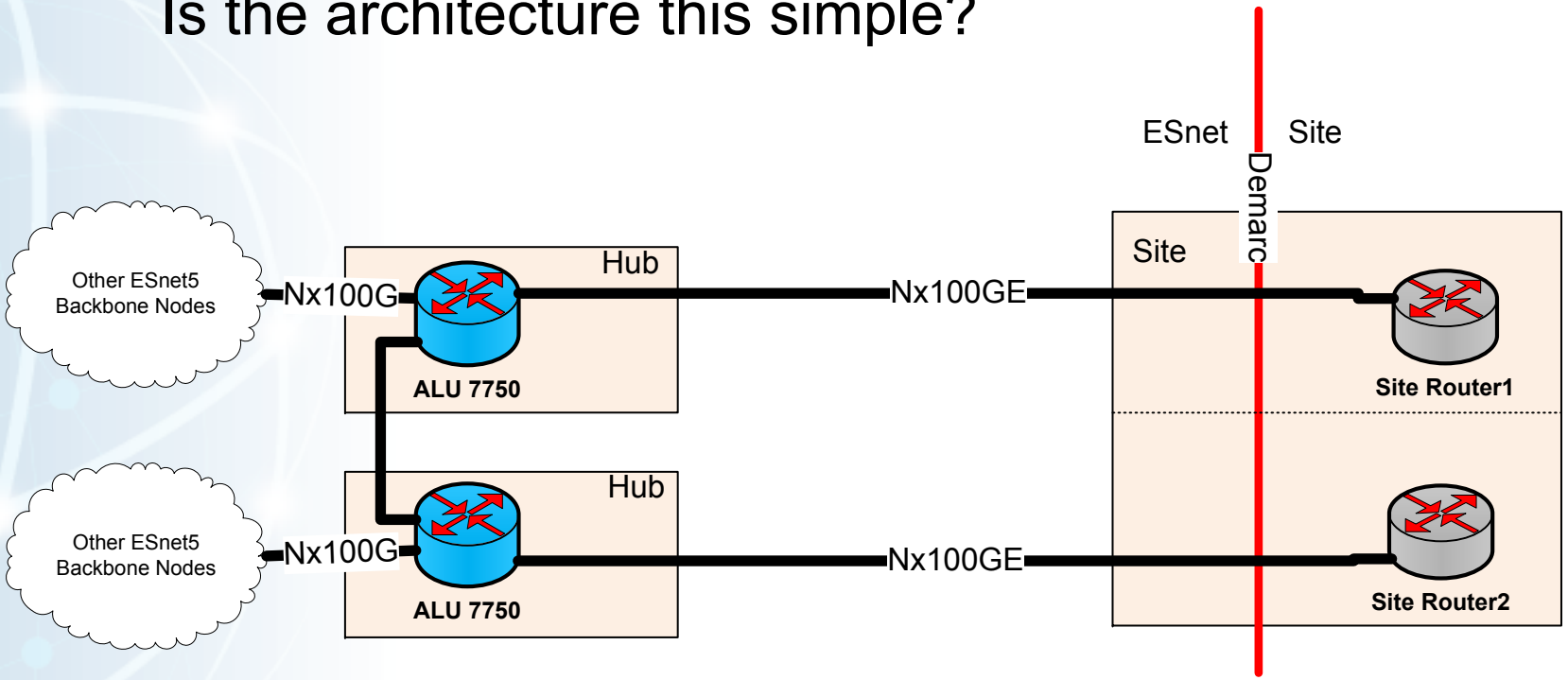


Which is better?



Site Connections beyond 200G

Is the architecture this simple?



Possible Additions to Network Services Portfolio

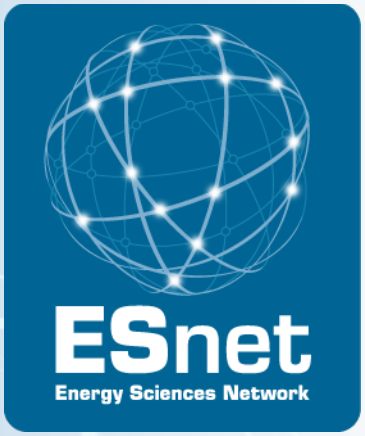
- Looking at additional types of overlay networks
  - Layer2 VPNs
    - Point to Point
    - Multipoint
  - Layer3 VPN's
    - IPV4 and IPV6
  - We haven't figured out which subset we need to support
  - Please let us know if you anticipate future requirements for overlay networks so we can take your needs into consideration

Additional NEG Updates

DNS DOS Rate Limiting

Improving site interactions

- Trying to figure out how to move from ad-hoc and event driven interactions to something that is more organized and intentional
- NEG Ambassador?



Questions?

Thanks!

Joe Metzger – metzger@es.net

http://www.es.net/

http://fasterdata.es.net/